Computing Optimal Commitments to Strategies and Outcome-Conditional Utility Transfers

Nathaniel Sauerberg
University of Texas–Austin
Austin, United States
njs@cs.utexas.edu

Caspar Oesterheld
Carnegie Mellon University
Pittsburgh, United States
oesterheld@cmu.edu

ABSTRACT

Prior work [14, 15] has studied the computational complexity of computing optimal strategies to commit to in Stackelberg or leadership games, where a leader commits to a strategy which is then observed by one or more followers. We extend this setting to one in which the leader can additionally commit to outcome-conditional utility transfers. In this setting, we characterize the computational complexity of finding optimal commitments for normal-form and Bayesian games. We find a mix of polynomial time algorithms and NP-hardness results. Then, we allow the leader to additionally commit to a signaling scheme based on her action, inducing a correlated equilibrium. In this variant, optimal commitments can be computed efficiently for arbitrarily many players.

KEYWORDS

Stackelberg Games; Leadership Games; Commitment Games; Computational Complexity; Transferable Utility; NP-Hardness; Linear Programming; Bayesian Games

ACM Reference Format:

Nathaniel Sauerberg and Caspar Oesterheld. 2024. Computing Optimal Commitments to Strategies and Outcome-Conditional Utility Transfers. In *Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), Auckland, New Zealand, May 6 – 10, 2024*, IFAAMAS, 10 pages.

1 INTRODUCTION

Since the early days of the field, game theorists have studied models of commitments and sequential play. Perhaps the first was von Stackelberg [44], who introduced his now-eponymous model to argue that a firm with the ability to commit to an action (a quantity of production) benefits from doing so in the duopoly competition model of Cournot [17]. Commitment to pure and mixed strategies were invoked by von Neumann and Morgenstern [43] to motivate the min-max and max-min values of zero-sum games. von Stengel and Zamir [45] study the payoff implications of commitment to mixed strategies, comparing the resulting payoffs to those in the Nash and correlated equilibria in the simultaneous versions of the game. The first to study the computational complexity of computing optimal commitments were Conitzer and Sandholm [15], who characterized the complexity of computing optimal pure and mixed strategies to commit to in normal-form and Bayesian games. Conitzer and Korzhyk [14] introduced the notion of commitment to correlated strategies and showed that finding optimal such commitments is tractable even with arbitrarily many players.

More recently, a number of authors in game theory and multi-agent learning have studied the use of voluntary utility transfers and commitments to pay money [2, 12, 20, 26, 31, 35, 37, 47–49]. We are specifically interested in outcome-conditional payments, such as committing to pay a particular agent if they take a particular action. We see two primary motivations to study such commitments to payments. The first is descriptive: Commitments to pay others are a common feature of human economies. For example, companies pay employees for performing tasks. The second is normative: It may be good if commitments to payments were available to and used by agents, as commitments to payments often allow for cooperative equilibria. For example, in the Prisoner's Dilemma, one player may pay the other to cooperate [12, 48].

In this paper, we combine these two lines of work and study games in which players can commit to both strategies and outcome-conditional utility transfers. To our knowledge, only Gupta and Schewe [26] have considered such a setting before. (See Section 1.2 for a discussion of their contributions.)

In the most basic case, we take a given normal-form game (the "base game") and imagine that Player 1 (the "leader") makes a commitment as follows: As in Stackelberg games, she commits to taking a specific action or mixture over actions in the base game. In addition, she can commit to transfer utility to other players based on the outcome of the game. Specifically, she commits to a payment function $P : A \rightarrow \mathbb{R}^{n-1}_{\geq 0}$, which we index $i = 2 \dots n$, where $P_i(a)$ is the amount of utility the leader will transfer to Player i if outcome a obtains.

For example, consider the game in Table 1. Since Bottom is a strictly dominant strategy for Player 1, the only Nash equilibrium is (Bottom, Middle), resulting in payoffs of $(0, 2)$. However, in our setting, the leader can commit to play the mixture $(1/3, 2/3)$ over (Top, Bottom) and to the payment function P where $P(B, L) = 1$ and $P(a) = 0$ for all $a \neq (B, L)$, essentially promising to transfer 1 utility to the follower if (Bottom, Left) is played.

Then, the follower's expected utility for playing Right is $2/3$, as is his utility for playing Left since the their payoffs after payments

Table 1: A game in which commitment to strategies and payments is beneficial to the leader.

	Left	Middle	Right
Top	−1, 0	−1, −2	−1, 0
Bottom	2, 0	0, 2	0, 1

Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), N. Alechina, V. Dignum, M. Dastani, J.S. Sichman (eds.), May 6 – 10, 2024, Auckland, New Zealand.

are the same. His utility for playing Middle is $2 \cdot 2/3 - 2 \cdot 1/3 = 2/3$. Hence, the follower is indifferent between each of his actions and tie-breaks in favor of the leader, playing Left for a leader payoff of $-1 \cdot 1/3 + (2-1) \cdot 2/3 = 1/3$. (Of course, the leader could also pay slightly more to strictly incentivize Left.)

It turns out that both commitment to actions and payments are necessary for the leader to achieve positive utility in this game. Intuitively, the leader cannot incentivize Left over Right without payments and cannot incentivize Left over Middle without commitment to (mixtures over) actions. For a more detailed analysis, see Proposition A.1 in the version at arXiv:2402.06626.

Settings involving commitment to both strategies and payments are ubiquitous in human interactions. For example, governments often pass laws (commitments) that determine both systems of subsidies (payments) as well as government policies that influence the structure of economic interactions (actions). These actions include what the government buys and sells, what services and entitlements it provides to citizens, what infrastructure it maintains, and how it regulates sectors.

Another example is companies' interactions with their employees. Here, companies set both general policies like rules, priorities, company hierarchy, etc. (commitment to strategies) and compensation policies like pay scales and bonus structure (commitment to payments), to which employees then respond. We might further imagine that leaders of subdivisions respond to the company's commitments by in turn setting general and compensation policies within their subdivisions, thus resulting in a multi-step game where players make commitments in sequence.

Both pure and mixed commitment are quite natural and common. In some cases, the followers observe the specific base game action the leader takes before making their decisions. In such cases, even if the leader does randomize, the followers would observe the specific draw from the distribution before making their choices, and so such randomization would not be helpful. In other cases, the followers can observe the leader's long-run distribution over actions before choosing their own actions, but cannot observe the leader's realized action in their specific instance.

1.1 Contributions

In this paper, we ask: How can we compute optimal commitments in such settings? The answer depends on various features of the setting, such as the number of players, whether only the first or also later players can commit, which players (if any) have Bayesian types and whether correlation devices are available. We give an overview of our results here.

In Section 3, we study a setting without private information and without correlation devices. (See Table 2 for a summary of the following results.)

- We show that in two-player games, optimal pure commitments can be found efficiently with dynamic programming (Theorem 3.1) and optimal mixed commitments can be found efficiently with linear programming (LP) (Theorem 3.2)
- We show that if there are more than two players and only the first commits, the optimal commitment is NP-complete to find (Theorem 3.3).
- We show that in the case of three players in which the players commit in sequence, the pure commitment case can be solved efficiently with LP (Theorem 3.4) while the mixed commitment case is NP-hard (Theorem 3.5).

In Section 4, we study a setting without private information but where correlation devices are available. Specifically, the leader can construct an arbitrary signaling scheme (as in [29]) that depends on her realized action. We show that in this setting, optimal mixed (Theorem 4.1) and pure (Theorem 4.2) commitments can both be computed efficiently with linear programming. (See Table 3 for a summary of these results.)

In Section 5, we extend the previous settings to Bayesian games, i.e., games where players have private information about their own preferences. (See Table 4 for a summary of the following results.)

- We show that in two-player games where only the follower has Bayesian types, the optimal commitment is NP-hard to compute under standard complexity-theoretic assumptions, regardless of the availability of correlation devices (Corollary 5.1.1).
- We show that in two-player games where only the leader has Bayesian types and correlation devices are not available, computing the optimal pure commitment is NP-hard (Theorem 5.2), while the optimal mixed commitment can be computed efficiently with LP (Theorem 5.3).
- We show that in n-player games where only the leader has Bayesian types and correlation devices are available, optimal mixed commitments can be computed in polynomial time with LP (Theorem 5.4).

Throughout the paper we give only sketches of our proofs. The full proofs can be found in the version at arXiv:2402.06626.

1.2 Related Work

The work of Gupta and Schewe [26] is most closely related to ours – they essentially consider the two-player, mixed-commitment version of our setting. Their focus is on comparing the cases where the follower tiebreaks for and against the leader and characterizing the properties of optimal commitments in more detail. They also provide an alternative proof of our Theorem 3.2.

As discussed in the introduction, von Stengel and Zamir [46] study a two-player setting with mixed commitment over actions, but without commitment to payments. They characterize the range of possible payoffs for the players under optimal commitment and compare them to the payoffs from other ways of playing the game, such as playing a Nash or correlated equilibrium of the simultaneous-move game, or playing a commitment game with leadership roles reversed. Conitzer and Sandholm [15] were the first to study the computational complexity of computing optimal strategies in Stackelberg games. They consider both pure and mixed commitment in both Bayesian and normal-form games. Follow up work [14] introduced the idea of committing to *correlated* strategies. Our present paper essentially extends the latter work to Bayesian games and extends both papers to allow the leader to commit to payments. Other works consider computing optimal commitments in extensive form [33] and Markov games [34]. Further variations on the setting include mixed commitments to which multiple followers must respond with pure strategies [13], and games with multiple leaders who commit simultaneously [8].

Another line of related work is that on k-implementation [20, 37], in which an outside "interested party" influences a normal-form game by committing to outcome-conditional payments, as in our setting. This is similar to our single commitment setting (Section 3.2), but the rationality assumptions of k-implementation allow the followers to play any undominated strategy, while we make the stronger assumption that the followers play according to a Nash or correlated equilibrium. Anderson et al. [2] apply the k-implementation framework to games the interested party plays in herself. They essentially characterize commitments to payments that optimize the leader's pure security level in the induced game. In reinforcement learning, settings like policy teaching [e.g. 50] and reward poisoning [e.g. 51] also study an outside party's ability to influence the agent's behavior via limited control over their reward function.

Payments have also been considered in the context of multi-agent reinforcement learning (MARL). Christoffersen et al. [12] also consider outcome-conditional payments. However, their setting has one player propose payment contracts for everyone, leaving the remaining players merely with the decision of whether to accept or reject the proposed set of payments. In contrast, in our setting, the leader unilaterally commits to payments but only for herself. Sodomka et al. [40] consider binding commitments to joint action profiles and side payments in two-player games, while other lines of work have considered (commitments to) sharing a fraction of one's reward *without* conditioning on the outcome [31, 48, 49] and even entirely unconditional gifting of reward, i.e., the transfer of some constant amount of money [35, 47].

The principal–agent problem literature in economics, also called contract theory, [e.g., 6, 16, 18, 23, 25, 32, 36, 42] also considers settings in which one player (the principal) can pay another player (the agent) to incentivize that player to take an action which induces a distribution over "outcomes", which the principal cares about. However, in this literature there is generally no notion of the principal taking actions herself, and hence no analog of commitment to actions. A key challenge is that the agent's action is generally unobservable except through its stochastic influence on the outcome, a difficulty our setting avoids. The principal-agent literature also generally assumes substantially more structure than the general normal-form games we consider. There is typically a single agent who maximizes their payment minus the cost of their action (though they are not always risk neutral). The rare settings with multiple agents typically have simple game structures, such as [6, 25, 39] in the principal-expert sub-literature and [4, 22], where the agents make binary decisions about whether to act and the outcome is either the AND or OR of these. Computational questions like those we consider are not typically the focus in principal-agent settings, though see [5, 21], as well as [1, 9, 27] for more recent work focused on settings with Bayesian types (which we also consider).

Another economic setting, mechanism design, relies implicitly on the commitment ability of the designer and frequently assumes transferable utility, but assumes the designer has more expansive abilities to control the rules of interaction than we consider here.

One alternative to payments is outcome-conditional commitments to burn utility [38, 41]. For example, Moulin [38] characterizes two-player normal-form games in which the payoffs in completely mixed equilibria can be improved by commitments to burn utility. The benefits of burning utility here come from changing one's own incentives and hence the game's Nash equilibrium, something which doesn't apply to our setting because the leader commits to her actions.

2 PRELIMINARIES

2.1 Normal-Form Games

An n-player normal-form game G is a pair (A, u), where $A = A_1 \times A_2 \times \cdots \times A_n$ is a non-empty, finite set of pure action profiles (or outcomes) and $u : A \to \mathbb{R}^n$ is a utility function mapping each outcome to a vector containing utilities for each player. Player i has action space A_i and utility function u_i. For 2-player games, we will sometimes denote the action space $A \times B$.

Let $\Delta(X)$ denote the set of probability distributions over a finite set X. In a normal-form game, a strategy for Player i is some $\sigma_i \in \Delta(A_i)$ and a strategy profile σ is a vector of strategies for each player. A strategy is *pure* if it plays one action with certainty. We'll use $-i$ to refer to the set of players besides Player i, so for example $\sigma_{-i} \in \Delta(A_{-i})$ and $\sigma = (\sigma_i, \sigma_{-i})$. For convenience, we'll define utilities over strategy profiles u in the obvious way: $u_i(\sigma) = \sum_{a\in A} \sigma(a)u_i(a)$, as we make the standard assumption that all players are risk-neutral expected utility maximizers.

An action a_i is a **best response** to σ_{-i} if $u_i(a_i, \sigma_{-i}) \geq u_i(a'_i, \sigma_{-i})$ for all $a'_i \in A_i$. A strategy σ_i is a best response to σ_{-i} if all supported actions are best responses. A **Nash equilibrium** (NE) of a game G is a strategy profile σ in which each σ_i is a best response to σ_{-i}. We use $\mathrm{NE}(G)$ to denote the set of Nash equilibria of a game G. A **correlated equilibrium** (CE) [3] of a game G is a distribution over outcomes $D \in \Delta(A)$ in which for each player i and each a_i with $D(a_i) > 0$, a_i is a best response to $D(\cdot \mid a_i) \in \Delta(A_{-i})$.

2.2 Commitment Games with Payments

We study games in which a distinguished player (the "leader") has the ability to modify a normal-form game by making a two-part commitment before it is played. First, she commits to taking a specific action or mixture over actions in the base game. Second, she can modify the payoff matrix by promising to transfer utility to another player whenever a certain outcome obtains. The leader's commitment is perfectly observed by all players.

Formally, a commitment is a pair (σ_1, P), where σ_1 is a strategy in the base game $G = (A, u)$ and $P : A \to \mathbb{R}^{n-1}_{\geq 0}$ is a payment function indexed $\{2, \ldots, n\}$ that intuitively means the leader commits to transfer $P_i(a)$ utility to Player i whenever outcome a obtains. A commitment induces a new normal-form game $G[\sigma_1, P]$ with $n-1$ players $\{2, \ldots, n\}$, strategy space $A_{-1} = A_2 \times A_3 \times \cdots \times A_n$, and utility functions $v_i^P(a_{-1}) = \sum_{a_1\in A_1} \sigma_1(a_1)[u_i(a) + P_i(a)]$. We'll often drop the superscript P from v_i when it's clear from context, and we'll refer to players $i \geq 2$ as the followers. While not a player in $G[\sigma_1, P]$, the leader still has utility function $v_1^P = u_1(a) - \sum_{i\geq 2} P_i(a)$ over the outcomes. As with utility functions, we'll extend the domain of P to include distributions over outcomes, i.e. $P(\sigma) = \sum_{a\in A} \sigma(a)P(a)$.

We'll sometimes require σ_1 to be a pure action in G, which we refer to as the **pure commitment** case, and we'll sometimes allow σ_1 to be a mixture over actions in G, which we refer to as the **mixed commitment** case.

When $n = 2$, $G[\sigma_1, P]$ is a single-player game so the follower simply takes an action maximizing his expected utility. We assume he tiebreaks in favor of the leader, so the outcome of any game is clear. With more than two players, however, it's not clear what happens after the leader commits. We study two variants.

In the **single commitment** setting, the followers play $G[\sigma_1, P]$ simultaneously as a typical normal-form game. This game may have multiple equilibria, so we'll assume the followers play the best Nash equilibrium for the leader. Hence, the leader's utility for a commitment (σ_1, P_1) is well defined as

$$\max_{\sigma_{-1} \in NE(G[\sigma_1, P])} u_1(\sigma_1, \sigma_{-1}) - \sum_j P_{1,j}(\sigma_1, \sigma_{-1}).$$

In the **sequential commitment** case, Player 2 becomes the leader of the (sequential) commitment game with base game $G[\sigma_1, P]$, which he plays optimally.[1] We assume that each player i has a lexicographic tiebreaking rule: They first maximize their own utility, then that of player $i - 1$, then $i - 2$, etc.[2] (We won't need to specify tiebreaking beyond this.) Since the leader's utility for a commitment is well-defined in two-player games, in the $n \geq 3$ sequential commitment case it's well-defined recursively. The leader's optimal commitments are the (σ_1, P_1) maximizing the leader's expected utility over the outcome of the sequential commitment game $G[\sigma_1, P_1]$.

3 NASH EQUILIBRIA

We begin by considering the case where the leader does not have access to a correlation device and there are no private types. We start with the simplest setting, with only two players. Note that with mixed commitment, the leader can always achieve at least the utility of the best-for-her Nash equilibrium of the base game, simply by committing to her strategy in that Nash equilibrium and zero payments. In contrast, the leader's utility under her optimal pure commitment may or may not exceed her utility in a Nash equilibrium of the base game. For proofs of these claims, see Appendix A in the version of this paper at arXiv:2402.06626.

3.1 Two Players

We first consider the two-player case where the leader commits to payments and pure actions. In this setting, the optimal commitments can be computed efficiently using dynamic programming.

Theorem 3.1. *In a two-player game, the leader's optimal commitment to a payment function and a pure action can be computed in polynomial time.*

Proof Sketch. To incentivize the follower to play a_2 after committing to play a_1 herself, the leader would need to pay him exactly $\max_{a_2'} u_2(a_1, a_2') - u_2(a_1, a_2)$. Therefore, we can easily compute the leader's maximum utility for implementing any outcome (a_1, a_2) and then maximize over them. □

Optimal mixed commitments in two-player games can also be computed in polynomial time. This result has already been shown [26, Corollary 11], but we include it here for completeness.

Theorem 3.2. *In a two-player game, the leader's optimal commitment to a payment function and a mixture over actions can be computed in polynomial time.*

Proof Sketch. For each follower action a_2, we construct an LP to compute the leader's optimal commitment that incentivizes a_2. There are variables corresponding to the follower's expected payment when a_2 is played and to the leader's probabilities of playing each action. Constraints ensure a_2 gives the follower at least as much utility as any other action. We can then simply maximize over the $|A_2|$ LPs to find the leader's optimal commitment. This is an extension of Conitzer and Sandholm [15]'s approach from the version of the setting without payments. □

3.2 More than Two Players, Single Commitment

We now consider games with $n \geq 3$ players, beginning with the single commitment case. Recall that in this setting, the leader makes a commitment (σ_1, P_1) and the followers are assumed to play the best Nash equilibrium for the leader of the induced game $G[\sigma_1, P_1]$.

We give a strong negative complexity result: Even if the leader has only a single action, computing the optimal commitment for her to make is NP-hard, even in a 3-player game. This immediately shows that both the pure and mixed commitment variants are NP-hard. Our proof is via reduction from the following problem:

Definition (Balanced Complete Bipartite Subgraph)**.** Given bipartite graph $G = (V, E)$ partitioned into partite sets V_1 and V_2 (s.t. E consists only of edges between V_1 and V_2) and a natural number k, decide whether there exist subsets $V_1' \subseteq V_1$ and $V_2' \subseteq V_2$ s.t. V_1' and V_2' each have (at least) k elements and there is an edge from each vertex in V_1' to each vertex in V_2'.

Balanced Complete Bipartite Subgraph is NP-complete [28, page 446]. It is sometimes called the Balanced Biclique Problem.

Theorem 3.3. *Consider an n-player game in which the leader commits to payments and a (mixture over) actions and then the remaining players play the best Nash equilibrium for the leader of the induced normal-form game. Computing the leader's optimal commitment is NP-hard, even for $n = 3$ players and for games in which the leader has only a single action.*

Proof Sketch. We give a reduction from Balanced Complete Bipartite Subgraph. The leader has only a single action, and roughly speaking, we design the game such that she cannot benefit from committing to payments. Players 2 and 3 have strategy spaces corresponding to the vertices of the graph, and a strategy profile is an equilibrium giving the leader high utility if and only if it corresponds to balanced complete bipartite subgraph of size k. Intuitively, this is because the followers benefit from playing adjacent vertices in their respective partite sets but cannot play any single vertex with probability more than $1/k$ without making themselves exploitable by the other follower. Our reduction is inspired

[1] In other words, we consider the subgame perfect Nash equilibria of our sequential commitment setting, as is typical in Stackelberg games without payments [e.g., 15, 45].
[2] We make this assumption to guarantee the existence of optimal commitments. For instance, if Player 2 broke ties against Player 1, Player 1 would be incentivized to give Player 2 a small additional payment ε so that Player 2 would strictly prefer the action benefiting Player 1, but no fixed ε could be optimal. In mechanism design, it is typically assumed that the agents break ties in favor of the leader for the same reason.

by the reduction of Gilboa and Zemel [24] from Clique to Best Nash (deciding whether there exists a Nash equilibrium giving a certain player at least k utility), but is simpler and reduces between different problems. □

3.3 More than Two Players, Sequential Commitment

We now consider the case where the players commit sequentially. That is, first Player 1 commits to (σ_1, P_1), then Player 2 commits to (σ_2, P_2), and so on. As noted before, we assume that each player i tiebreaks in favor of $i-1$, then $i-2$, etc. First, we give an efficient algorithm for the $n = 3$ player pure commitment case, leaving the $n > 3$ player case as an open problem.

Theorem 3.4. *In a three-player game in which the players commit sequentially to payment functions and pure actions, the leader's optimal commitment can be computed in polynomial time.*

Proof Sketch. We construct an LP that computes, for any given outcome a, an optimal leader commitment that implements a. To do so, we first show that, intuitively, it's optimal for the leader to pay Player 3 some extremely large amount to minimize Player 2's utility if Player 2 doesn't play a_2. That is, for each action $a_2' \neq a_2$, the leader commits to make a large payment to Player 3 when (a_1, a_2', a_3^*) is played, where $a_3^* = \min_{a_3'} u_2(a_1, a_2', a_3')$. Since the LP has incentive constraints that ensure a will be played, these large off-equilibrium payments will never actually need to be made. The LP has 3 variables corresponding to the on-equilibrium payments (including from Player 2 to 3) and incentive constraints ensuring that Player 2 prefers to implement a rather than either deviate himself or implement a deviation by Player 3. As usual, we then maximize over the $|A|$ possible outcomes. □

When players commit sequentially to *mixtures* over actions, computing the optimal commitment is NP-hard (as in the case without payments [15]), even with only three players. We show this via reduction from the following problem.

Problem 3.1 (Balanced Vertex Cover)**.** Given a graph $G = (V, E)$, decide whether there exists a subset of vertices $S \subseteq V$ of size at most $|V|/2$ such that, for all edges $e = (v_1, v_2) \in E$, at least one of v_1 and v_2 is in S.

Given a subset of vertices S, we say an $e = (v_1, v_2)$ is "covered" if at least one of v_1 and v_2 is in S, and otherwise we say it is "uncovered". If S covers all edges in G, it is called a *cover* of G, or a K-cover if it has cardinality K.

Balanced Vertex Cover is the special case of Vertex Cover in which the size of the requested cover is $|V|/2$. Vertex Cover was one of Karp's original 21 NP-complete problems [30]. Conitzer and Sandholm [15] define Balanced Vertex Cover and show it remains NP-complete. Intuitively, an arbitrary Vertex Cover instance can be reduced to a balanced instance by either adding isolated vertices while $K > |V|/2$ or by adding isolated triangles and increasing K by 2 per triangle while $K < |V|/2$.

Theorem 3.5. *In an n-player game in which the players commit sequentially to payment functions and mixtures over actions, computing the leader's optimal commitment is NP-hard, even for $n = 3$ players.*

Table 2: Overview of Results from Section 3

	Pure Commitment	Mixed Commitment
$n = 2$	$\Theta(\|A\|)$ time via DP (Theorem 3.1)	1 LP-solve per follower action (Theorem 3.2)
$n = 3$, single	NP-hard, even with only one leader action (Theorem 3.3)	
$n = 3$, sequential	1 LP-solve per action profile (Theorem 3.4)	NP-hard (Theorem 3.5)

Proof Sketch. We reduce from Problem 3.1, Balanced Vertex Cover. We construct a game in which the first two players each have an action for each vertex and play "cooperatively" because they share the same utility function. They can achieve high utility if and only if the first mixes uniformly over vertices corresponding to a balanced vertex cover and the second mixes uniformly over its complement. The third player has actions that "exploit" the first two if they don't play a balanced vertex cover, meaning the first player covers every edge and together they play every vertex with high enough probability. If no such exploit would succeed, the third player will instead play a different action giving the first two players high utility. This approach adapts the reduction of Conitzer and Sandholm [15, Theorem 4] to the version of the setting without payments, modifying their construction slightly so that the ability to commit to payments cannot benefit the first two players. □

4 CORRELATED EQUILIBRIA

We will now consider allowing the leader to commit to a signaling scheme (as in [29]) along with her payments and actions, inducing a correlated equilibrium. Specifically, the leader picks a probability distribution $D \in \Delta(A)$ and a payment function $P_j : A_j \to \mathbb{R}_{\geq 0}$ for each other player j. Intuitively, $P_j(a_j)$ is the amount paid to Player j for following a recommendation to take action a_j. After committing to D, the leader privately draws an action profile $a \sim D$, sends each player $j \geq 2$ the private recommendation a_j, and plays a_1 herself. Finally, the followers simultaneously choose their actions in the base game, making no commitments of their own. Note that the leader's commitment (D, P) includes a commitment to actions since she commits to play according to the draw from D.

We give a revelation principle style result (Lemma C.1 in the version at arXiv:2402.06626) showing that any leader commitment is equivalent (in terms of the distribution over outcomes and payoffs) to an **incentive compatible** commitment in which all followers are incentivized to follow their recommended actions. Hence, we will assume without loss of generality that the leader's commitment is incentive compatible. We also show that allowing the leader to send arbitrary messages makes no difference compared to allowing only action recommendations and that allowing payments to depend on the full outcome and recommendation profile is no more powerful than simply paying players for following their recommended action. (Again, see Lemma C.1 for details.)

Formally, a commitment (D, P) is **incentive compatible** if, for any player j and any action a_j with $D(a_j) > 0$, player j is always (weakly) incentivized to follow a recommendation to play a_j, assuming all other players follow their recommendations. That is, for all players j, actions a_j with $D(a_j) > 0$, and all $a'_j \neq a_j$,

$$\sum_{a_{-j}} D(a_{-j}|a_j)u_j(a_j, a_{-j}) + P_j(a_j) \geq \sum_{a_{-j}} D(a_{-j}|a_j)u_j(a'_j, a_{-j}).$$

We call these the incentive constraints for player j. Hence, the problem of computing the leader's optimal commitment is equivalent to that of finding a utility maximizing distribution D over outcomes and payment function P, subject to the incentive constraints. (Note that there is no incentive constraint for the leader since she commits to her action, though she is sometimes constrained to play a pure action.)

It turns out that an optimal commitment to a signaling scheme can be found in polynomial time regardless of whether the leader can commit to mixtures or only to pure actions, as the following the results show.

Theorem 4.1. *In an n-player game, the leader's optimal commitment to a payment function, mixture over actions, and signaling scheme can be computed in polynomial time.*

Proof Sketch. We construct a linear program which computes the leader's optimal incentive-compatible commitment. Such commitments are optimal by Lemma C.1. Our LP is similar to the feasibility LP for a correlated equilibrium except that it incorporates payments. In addition, the leader's incentive constraints are dropped because she commits to follow her action recommendation and the objective is the leader's expected utility (after payments).

The LP has variables p_a for each action profile a representing the that action profile is recommended and variables $t_i(a_i)$ representing the payment from the leader to each follower i when an recommendation to play action a_i is followed. Constraints ensure that, for each follower and each possible action recommendation, the follower's expected utility for following the recommendation (including payments) is at least their expected utility for any other action. This expectation is taken over the player's uncertainty over the other players' action recommendations given their own. (The constraints are designed such that those corresponding to actions that are never recommended are trivially satisfied.)

Our approach is similar to and inspired by that of Conitzer and Korzhyk [14] for computing optimal correlated distributions to commit to without payments. □

Theorem 4.2. *In an n-player game, the leader's optimal commitment to a payment function, pure action, and signaling scheme can be computed in polynomial time.*

Proof Sketch. For any leader action a_1, we can add a constraint to the LP from Theorem 4.1 to require the leader to play the pure strategy a_1, which allows us to compute the leader's optimal commitment and corresponding utility when committing to a_1. Solving such an LP for all actions a_1 and maximizing over their values gives the leader's optimal commitment overall. □

5 BAYESIAN GAMES

We now consider settings with Bayesian, rather than normal-form, base games. In normal-form games, the agents' utility functions are common knowledge. Bayesian games, in contrast, model agents with private information about their own preferences. In Bayesian games, each player i has a set of types Θ_i and for each type θ_i, a utility function $u_i^{\theta_i} : A \to \mathbb{R}$. In games in which only one player has multiple types, we often drop the subscript i from θ and Θ.

Table 3: Overview of Results from Section 4

	Pure Commitment	Mixed Commitment
Any n	1 LP-solve (Theorem 4.1)	1 LP-solve per leader action (Theorem 4.2)

The distribution $\pi_i \in \Delta(\Theta_i)$ over each player's type is common knowledge before the game begins (*ex ante*). Then, *ex interim*, each player i learns their own realized type θ_i and chooses an action. Hence, a strategy for Player i in a Bayesian game is a mapping $\sigma_i : \Theta_i \to \Delta(A_i)$ which specifies a mixture over actions for each potential type. An action a_i is a best response for type θ_i if $\mathbb{E}_{\theta_{-i}|\theta_i}[u_i(a_i, \sigma_{-i}(\theta_{-i}))] \geq \mathbb{E}_{\theta_{-i}|\theta_i}[u_i(a'_i, \sigma_{-i}(\theta_{-i}))]$ for all $a'_i \in A_i$, and a strategy σ_i is a best response to σ_{-i} if all actions supported in each $\sigma_i(\theta_i)$ are best responses for θ_i. A strategy profile σ is a **Bayes-Nash equilibrium** (BNE) if each σ_i is a best response to σ_{-i}.

It turns out that leader and follower types introduce different considerations and have different impacts on the complexity of computing optimal commitments, so we'll consider the two cases separately. (When we refer to a setting with leader types, we mean the special case of Bayesian games in which the follower has only one type, and vice versa.) In addition, we have all the same variations in the Bayesian setting as we did without private information: pure vs mixed commitment, sequential vs single commitment, and whether or not the leader is able to commit to a signaling scheme.

It's important to note that the commitments are to potentially different (mixtures over) actions for each type. In particular, in the pure commitment case, each type must play a pure action, but different types can play different actions. If the leader's actions couldn't depend on her type, this setting would be equivalent to one without leader types in which the leader's utility was equal to her expected utility over types. We also make the natural assumptions that the leader's payment cannot depend on the follower's type and that the followers don't gain any information about the leader's realized type before taking their action (except insofar as their action recommendations are correlated with the leader's type).

5.1 Follower Types

First, consider the case with $n = 2$ agents where the leader has only one type but has Bayesian uncertainty over the follower's type. Without payments, computing the optimal pure action for the leader to commit to is quite simple: for each possible leader action, one can compute each follower type's best response and hence the leader's expected utility [15, Theorem 6].

However, the problem becomes difficult if the leader can additionally commit to payments. We show a surprising connection to auction theory, reducing to our present setting from problem of finding a revenue-maximizing item pricing for m items and a single unit-demand buyer.

A unit-demand buyer is one who essentially "wants" at most one item, i.e. considers the items perfect substitutes. Formally, a unit-demand buyer has a value v_i for each item i and their utility for receiving a set of items S is the maximum over their values for items in the set $\max_{i \in S} v_i$. An item pricing (or posted pricing) offers each item to the buyer at a take-it-or-leave-it price and allows the buyer to purchase whatever items they want, as at a typical retail store. Formally, this is a vector r of non-negative prices, one for each item. Therefore, given an item pricing, a unit-demand buyer will purchase a single item maximizing his value minus the price or, if all prices are greater than the items' values to the buyer, will purchase nothing.

Problem 5.1 (Unit Demand Item Pricing)**.** Consider a finite-support distribution D of value vectors $v \in \mathbb{R}^m_{\geq 0}$. For an item pricing vector $r \in \mathbb{R}^m_{\geq 0}$, let $i^*(v) \in \arg\max_i [v_i - r_i]$ be the most favorable item for a buyer with values v to buy, tiebreaking in favor of the item with highest value v_i. Let $A(r) = \{v \in \text{supp}(D) | v_{i^*(v)} - r_{i^*(v)} \geq 0\}$ be the set of values v for which the buyer buys an item. Decide whether there exists r such that the seller's revenue $\text{Rev}(r) = \sum_{v \in A(r)} [D(v) \cdot r_{i^*(v)}]$ is at least K.

Briest [7] gives a hardness result for a special case of Unit Demand Pricing they define, which we'll call Unit Demand Pricing for Uniform Budgets. The special case requires each value vector v to have some "budget" β_v such that all entries $v_i \in \{0, \beta_v\}$. They refer to it as the unit-demand min-buying (or envy-free) pricing problem with uniform budgets, economist's version. Specifically, they show Unit Demand Pricing for Uniform Budgets is NP-hard to approximate subpolynomially in the number of items unless NP is in bounded-error probabilistic sub-exponential time:

Theorem ([7], Theorem 5)**.** Unit Demand Pricing for Uniform Budgets is hard to approximate within $O(|G|^\varepsilon)$ for some $\varepsilon > 0$ if $\mathsf{NP} \nsubseteq \cap_{\delta > 0} \mathsf{BPTIME}(2^{O(n^\delta)})$.

The general problem of Unit Demand Item Pricing is widely believed to unconditionally be NP-hard [10], perhaps in part because the version where the buyer's values for items are independent rather that correlated is NP-hard [11].

We now give a reduction from general Unit Demand Item Pricing to our present setting of computing optimal pure action and payment commitments in 2-player Bayesian games. We'll therefore inherit the hardness result from the uniform budgets special case.

Theorem 5.1. Unit Demand Item Pricing is polynomial-time reducible to the problem of computing the leader's optimal payments and (mixtures over) actions in a two-player Bayesian game with follower types only. Further, it is reducible to instances in which the leader has only a single action.

Proof Sketch. We reduce an arbitrary instance U of Unit Demand Item Pricing to a 2-player Bayesian game G with a single leader type and single leader action. In G, the follower has actions b_i corresponding to each item in U and types θ_v corresponding to each value vector v in U. Each type θ_v occurs in G with the same probability as v in U. We construct payoffs such that taking action t_i corresponds to purchasing item i in u and the leader's payment function corresponds to an item pricing in U. Specifically, the leader receives some very large utility Z whenever the follower plays any action b_i, but the follower of type θ_v receives utility $-Z + v_i$ for playing b_i. Committing to a payment of $Z - r_i$ for action b_i then corresponds to setting a posted price of r_i: if b_i is played the leader gets utility $Z - (Z - r_i) = r_i$ and the follower gets utility $(-Z + v_i) + (Z - r_i) = v_i - r_i$. Because the follower in G plays the single action that maximizes his utility, his behavior is equivalent to a unit demand buyer in U, who purchases the single item which maximizes his utility. ◻

Corollary 5.1.1. In a two-player Bayesian game with follower types only, computing the leader's optimal commitment to a payment function and a (mixture over) actions is NP-hard, even if the leader has only a single action, assuming $\mathsf{NP} \nsubseteq \cap_{\delta > 0} \mathsf{BPTIME}(2^{O(n^\delta)})$.

Because our construction requires only a single leader action, it immediately shows the hardness of both the pure and mixed commitment versions of the problem. Likewise, with only $n = 2$ players a single leader action, there is clearly no possible correlation to be had, and so access to signaling devices makes no difference. Since this hardness result applies to any version of the setting with follower types, we now consider settings with leader types only.

5.2 Leader Types–Nash Equilibrium

When the leader has types, it's not immediately obvious whether to model the leader as already knowing her type at the time of commitment: Should the leader make her commitment *ex ante* or *ex interim*? However, if the leader with types commits *ex interim* with a realized type θ_i, the problem instance is equivalent to one *without* leader types in which the leader's utility function is $u_1^{\theta_i}$: Once the leader commits to her strategy, the follower(s) simply best respond and the leader's utility function (and hence her type and uncertainty over it) has no further impact on the game.[3]

Therefore, we'll assume the leader seeks to maximize her *ex ante* expected utility, where the expectation is over the randomness in her realized type (as well as the randomness in the game, of followers' types, etc.). In this setting, a commitment is a payment function and a mapping from types to actions. Note that it would make no difference if the leader could commit to different payment functions for different types: The follower's followers' actions and leader's *ex ante* expected utility depend only on the expected payment function (over leader types).

Unlike with follower types, the pure and mixed commitment settings with leader types differ in hardness. We'll begin with the former. Even for $n = 2$ players, the analog of the problem without payments is NP-hard [15, Theorem 5]. We now show this hardness continues to hold when the leader can commit to payments.

Theorem 5.2. In a two-player Bayesian game with leader types only, computing the leader's *ex ante* optimal commitment to a payment function and an action for each type is NP-hard.

Proof Sketch. We prove this via reduction from Vertex Cover. The leader has K types, each of which occur with equal probability, and an action a_v corresponding to each vertex v in the graph. We show that she achieves strictly positive utility if and only if she commits to a strategy where, for each vertex v in a K-cover of G, a_v is the action of one of her types. The follower has strategies b_e corresponding to each edge in $e \in E$ and an additional strategy b_0. All leader types get utility 1 if the follower plays b_0 and 0 otherwise.

[3]Except that we assume the follower(s) break ties in favor of the leader's realized type.

The follower's utility function is such that, if all edges are covered by some leader type, the follower will prefer to play b_0, resulting in a leader utility of 1. However, if there exists an "uncovered" edge e for which no leader type plays a vertex $v \in e$, the follower will prefer to play b_e rather than b_0, resulting in a leader utility of 0. The leader cannot pay the follower enough to incentivize him to play b_0 when an uncovered edge exists while still getting positive utility herself. Therefore, the leader can achieve strictly positive utility if and only if there exists a K-cover in the Vertex Cover instance.

Our reduction is the same as that of Conitzer and Sandholm [15] to the version of this setting without commitment to payments. □

In contrast to the pure commitment case, the mixed commitment case with leader types is tractable with $n = 2$ players.

Theorem 5.3. In a two-player Bayesian game with leader types only, the leader's *ex-ante* optimal commitment to a payment function and a mixture over actions for each type can be computed in polynomial time.

Proof Sketch. We extend our LP approach from Theorem 3.2, which considers the analogous setting without types. Our LP variables now encode a distribution over actions for each leader type and optimizes the leader's *ex ante* expected utility over her types, again subject to the constraint that the follower is incentivized to play a particular pure action. As usual, we then maximize over the $|B|$ follower actions. □

If the leader doesn't have access to a signaling device, the problem becomes hard for $n \geq 3$: Even without leader types, we've already given hardness results in Theorem 3.3 and Theorem 3.5 for both single and sequential mixed commitment, respectively.

5.3 Leader Types–Correlated Equilibrium

Hence, we turn to the case where the leader can commit to a signaling scheme. With signaling, the problem becomes tractable for any number of players n.

Theorem 5.4. In an n-player Bayesian game with leader types only, the leader's *ex ante* optimal commitment to a payment function, a mixture over actions for each type, and a signaling scheme can be computed in polynomial time.

Proof Sketch. We extend our linear programming approach from Theorem 4.1, which considers the analogous setting without types. Our LP variables now encode a different distribution over outcomes for each leader type, as well as payments. We optimize the leader's *ex ante* expected utility over her types, and the incentive constraints now hold in expectation over the leader's types. □

One may recall that for the normal-form case with access to signaling devices, we could extend the mixed commitment approach to the pure setting by simply having one LP for each leader action. Taking the same approach doesn't work with leader types, however, because commitments to pure actions are now functions from leader types to actions, rather than just single actions. Hence, we'd need one LP for each of the $\Omega(|A_1|^{|\Theta|})$ commitment functions rather than just one per leader action. Indeed, with pure commitment and leader types, we already have the hardness result from the $n = 2$ case (Theorem 5.2), in which signaling is not helpful.

Table 4: Summary of Results from Section 5

		Pure Commitment	**Mixed Commitment**
Follower Types	$n = 2$	NP-hard under standard complexity-theoretic assumptions (Corollary 5.1.1)	
Leader Types	$n = 2$	NP-hard (Theorem 5.2)	1 LP solve per follower action (Theorem 5.3)
	$n = 3$, no signaling	—-	NP-hard even without types (Theorem 3.3, Theorem 3.5)
	Any n, signaling	—-	1 LP solve (Theorem 5.4)

6 CONCLUSION

In this work, we've presented a model combining Stackelberg games with outcome-conditional utility transfers and analyzed the computational complexity of computing optimal commitments. We've varied our setting along several dimensions: whether the leader commits to mixed or pure actions, whether the leader can commit to a signaling scheme, whether the followers play simultaneously or make commitments sequentially, and whether players have private information about their payoffs. We've given a mixture of efficient algorithms for computing optimal commitments, primarily via linear programming, and NP-hardness results. For an overview of the specific results, we refer back to Tables 2, 3 and 4.

There are many open directions for future work. One natural idea is to extend the framework of commitments to both actions and payments to game representations besides normal form. Another is to vary the assumptions we make about the payments. For instance, one could consider settings where there are costs associated with commitment, akin to the legal costs of writing a contract. One could also consider restrictions on what the payments can depend on, for instance if only certain actions can be detected and thus have payments associated with them.

One could also consider weakening the strong, completely binding model of commitment we've studied in this work. For instance, agents might only be able to commit to certain aspects of actions, or they might be able to commit *against* taking certain pure actions without being able to commit to a mixture over the remaining actions (as in [19]). Examples of weaker levels of commitment are prevalent in economic and societal interactions. For instance, firms give press releases and may fear the public relations or stock price ramifications of reversing course. Similarly, they make investments that make the indicated actions very likely, but not entirely certain.

ACKNOWLEDGMENTS

We thank Vincent Conitzer for helpful discussions. Nathaniel Sauerberg acknowledges financial support from the Center on Long-Term Risk and ML Alignment and Theory Scholars. Caspar Oesterheld acknowledges financial support through grants from the Cooperative AI Foundation, Polaris Ventures (formerly the Center for Emerging Risk Research), and Jaan Tallinn's donor-advised fund at Founders Pledge. Additionally he is grateful for support by an FLI PhD Fellowship.

REFERENCES

[1] Tal Alon, Paul Dütting, and Inbal Talgam-Cohen. 2021. Contracts with Private Cost per Unit-of-Effort. In *Proceedings of the 22nd ACM Conference on Economics and Computation* (Budapest, Hungary) *(EC '21)*. Association for Computing Machinery, New York, NY, USA, 52–69. https://doi.org/10.1145/3465456.3467651

[2] Ashton Anderson, Yoav Shoham, and Alon Altman. 2010. Internal Implementation. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: Volume 1* (Toronto, Canada) *(AAMAS '10)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 191–198.

[3] Robert J Aumann. 1987. Correlated equilibrium as an expression of Bayesian rationality. *Econometrica: Journal of the Econometric Society* (1987), 1–18.

[4] Moshe Babaioff, Michal Feldman, Noam Nisan, and Eyal Winter. 2012. Combinatorial agency. *Journal of Economic Theory* 147, 3 (2012), 999–1034. https://doi.org/10.1016/j.jet.2012.01.010

[5] Moshe Babaioff and Eyal Winter. 2014. Contract Complexity. In *Proceedings of the Fifteenth ACM Conference on Economics and Computation* (Palo Alto, California, USA) *(EC '14)*. Association for Computing Machinery, New York, NY, USA, 911. https://doi.org/10.1145/2600057.2602852

[6] John M. Barron and Glen R. Waddell. 2003. Executive rank, pay and project selection. *J. Financ. Econ.* 67 (2003), 305–349.

[7] Patrick Briest. 2008. Uniform Budgets and the Envy-Free Pricing Problem. In *Proceedings of the 35th International Colloquium on Automata, Languages and Programming - Volume Part I* (Reykjavik, Iceland) *(ICALP '08)*. Springer-Verlag, Berlin, Heidelberg, 808–819. https://doi.org/10.1007/978-3-540-70575-8_66

[8] Matteo Castiglioni, Alberto Marchesi, and Nicola Gatti. 2019. Be a Leader or Become a Follower: The Strategy to Commit to with Multiple Leaders. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19*. International Joint Conferences on Artificial Intelligence Organization, 123–129. https://doi.org/10.24963/ijcai.2019/18

[9] Matteo Castiglioni, Alberto Marchesi, and Nicola Gatti. 2021. Bayesian Agency: Linear versus Tractable Contracts. In *Proceedings of the 22nd ACM Conference on Economics and Computation* (Budapest, Hungary) *(EC '21)*. Association for Computing Machinery, New York, NY, USA, 285–286. https://doi.org/10.1145/3465456.3467602

[10] Shuchi Chawla and Balasubramanian Sivan. 2014. Bayesian Algorithmic Mechanism Design. *SIGecom Exch.* 13, 1 (nov 2014), 5–49. https://doi.org/10.1145/2692375.2692378

[11] Xi Chen, Ilias Diakonikolas, Dimitris Paparas, Xiaorui Sun, and Mihalis Yannakakis. 2018. The complexity of optimal multidimensional pricing for a unit-demand buyer. *Games and Economic Behavior* 110 (2018), 139–164. https://doi.org/10.1016/j.geb.2018.03.016

[12] Phillip JK Christoffersen, Andreas A Haupt, and Dylan Hadfield-Menell. 2023. Get It in Writing: Formal Contracts Mitigate Social Dilemmas in Multi-Agent RL. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems*. IFAAMAS.

[13] Stefano Coniglio, Nicola Gatti, and Alberto Marchesi. 2020. Computing a pessimistic stackelberg equilibrium with multiple followers: The mixed-pure case. *Algorithmica* 82, 5 (2020), 1189–1238.

[14] Vincent Conitzer and Dmytro Korzhyk. 2011. Commitment to Correlated Strategies. *Proceedings of the AAAI Conference on Artificial Intelligence* 25, 1 (Aug. 2011), 632–637. https://doi.org/10.1609/aaai.v25i1.7875

[15] Vincent Conitzer and Tuomas Sandholm. 2006. Computing the Optimal Strategy to Commit To. In *Proceedings of the 7th ACM Conference on Electronic Commerce* (Ann Arbor, Michigan, USA) *(EC '06)*. Association for Computing Machinery, New York, NY, USA, 82–90. https://doi.org/10.1145/1134707.1134717

[16] John E Core, Jun Qian, et al. 2002. Project selection, production, uncertainty, and incentives. *Production, Uncertainty, and Incentives (January 2002)* (1 2002). https://doi.org/10.2139/ssrn.297461

[17] Antoine Augustin Cournot. 1838. *Recherches sur les principes mathématiques de la théorie des richesses*. L. Hachette.

[18] Joel S. Demski and David E. M. Sappington. 1987. Delegated Expertise. *J. Account. Res.* 25, 1 (1987), 68–89.

[19] Yuan Deng and Vincent Conitzer. 2017. Disarmament Games. *Proceedings of the AAAI Conference on Artificial Intelligence* 31, 1 (Feb. 2017). https://doi.org/10.1609/aaai.v31i1.10573

[20] Yuan Deng, Pingzhong Tang, and Shuran Zheng. 2016. Complexity and Algorithms of K-Implementation. In *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems* (Singapore, Singapore) *(AAMAS '16)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 5–13.

[21] Paul Dütting, Tim Roughgarden, and Inbal Talgam-Cohen. 2019. Simple versus Optimal Contracts. In *Proceedings of the 2019 ACM Conference on Economics and Computation* (Phoenix, AZ, USA) *(EC '19)*. Association for Computing Machinery, New York, NY, USA, 369–387. https://doi.org/10.1145/3328526.3329591

[22] Yuval Emek and Michal Feldman. 2012. Computing optimal contracts in combinatorial agencies. *Theoretical Computer Science* 452 (2012), 56–74. https://doi.org/10.1016/j.tcs.2012.05.018

[23] Eberhard Feess and Markus Walzl. 2004. Delegated expertise – when are good projects bad news? *Econ. Lett.* 82, 1 (1 2004), 77–82.

[24] Itzhak Gilboa and Eitan Zemel. 1989. Nash and correlated equilibria: Some complexity considerations. *Games and Economic Behavior* 1, 1 (1989), 80–93.

[25] Denis Gromb and David Martimort. 2007. Collusion and the organization of delegated expertise. *J. Econ. Theory* 137, 1 (11 2007), 271–299.

[26] Anshul Gupta and Sven Schewe. 2015. It Pays to Pay in Bi-Matrix Games: A Rational Explanation for Bribery. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems* (Istanbul, Turkey) *(AAMAS '15)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 1361–1369.

[27] Guru Guruganesh, Jon Schneider, and Joshua R. Wang. 2021. Contracts under Moral Hazard and Adverse Selection. In *Proceedings of the 22nd ACM Conference on Economics and Computation* (Budapest, Hungary) *(EC '21)*. Association for Computing Machinery, New York, NY, USA, 563–582. https://doi.org/10.1145/3465456.3467637

[28] David S Johnson. 1987. The NP-completeness column: An ongoing guide. *Journal of Algorithms* 8, 3 (1987), 438–448. https://doi.org/10.1016/0196-6774(87)90021-6

[29] Emir Kamenica and Matthew Gentzkow. 2011. Bayesian persuasion. *American Economic Review* 101, 6 (2011), 2590–2615.

[30] Richard M. Karp. 1972. *Reducibility among Combinatorial Problems*. Springer US, Boston, MA, 85–103. https://doi.org/10.1007/978-1-4684-2001-2_9

[31] Michael Kölle, Tim Matheis, Philipp Altmann, and Kyrill Schmid. 2023. Learning to Participate Through Trading of Reward Shares. In *Proceedings of the 15th International Conference on Agents and Artificial Intelligence - Volume 1: ICAART*. INSTICC, SciTePress, 355–362. https://doi.org/10.5220/0011781600003393

[32] Richard A. Lambert. 1986. Executive Effort and Selection of Risky Projects. *RAND J. Econ.* 17, 1 (1986), 77–88.

[33] Joshua Letchford and Vincent Conitzer. 2010. Computing Optimal Strategies to Commit to in Extensive-Form Games. In *Proceedings of the 11th ACM Conference on Electronic Commerce* (Cambridge, Massachusetts, USA) *(EC '10)*. Association for Computing Machinery, New York, NY, USA, 83–92. https://doi.org/10.1145/1807342.1807354

[34] Joshua Letchford, Liam MacDermed, Vincent Conitzer, Ronald Parr, and Charles Isbell. 2021. Computing Optimal Strategies to Commit to in Stochastic Games. *Proceedings of the AAAI Conference on Artificial Intelligence* 26, 1 (Sep. 2021), 1380–1386. https://doi.org/10.1609/aaai.v26i1.8267

[35] Andrei Lupu and Doina Precup. 2020. Gifting in Multi-Agent Reinforcement Learning. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems* (Auckland, New Zealand) *(AAMAS '20)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 789–797.

[36] James M Malcomson. 2009. Principal and Expert Agent. *The B.E. Journal of Theoretical Economics* 9, 1 (2009). https://doi.org/10.2202/1935-1704.1528

[37] Dov Monderer and Moshe Tennenholtz. 2004. k-Implementation. *Journal of Artificial Intelligence Research* 21 (2004), 37–62.

[38] Hervé Moulin. 1976. Cooperation in mixed equilibrium. *Mathematics of Operations Research* 1, 3 (1976), 273–286.

[39] Caspar Oesterheld and Vincent Conitzer. 2020. Minimum-regret contracts for principal-expert problems. In *Proceedings of the 16th Conference on Web and Internet Economics (WINE)*.

[40] Eric Sodomka, Elizabeth Hilliard, Michael Littman, and Amy Greenwald. 2013. Coco-Q: Learning in Stochastic Games with Side Payments. In *Proceedings of the 30th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 28)*, Sanjoy Dasgupta and David McAllester (Eds.). PMLR, Atlanta, Georgia, USA, 1471–1479. https://proceedings.mlr.press/v28/sodomka13.html

[41] Filipe Souza and Leandro Rêgo. 2012. *Mixed Equilibrium: When Burning Money is Rational*. MPRA Paper. University Library of Munich, Germany. https://EconPapers.repec.org/RePEc:pra:mprapa:43410

[42] Neal M. Stoughton. 1993. Moral Hazard and the Portfolio Management Problem. *J. Finance* 48, 5 (12 1993), 2009–2028.

[43] John von Neumann and Oskar Morgenstern. 1947. Theory of games and economic behavior, 2nd rev. (1947).

[44] Heinrich von Stackelberg. 1934. *Marktform und Gleichgewicht*. Springer.

[45] Bernhard von Stengel and Shmuel Zamir. 2004. *Leadership with commitment to mixed strategies*. Technical Report. Citeseer.

[46] Bernhard von Stengel and Shmuel Zamir. 2010. Leadership Games with Convex Strategy Sets. *Games and Economic Behavior* 69 (2010), 446–457.

[47] Woodrow Z. Wang, Mark Beliaev, Erdem Bıyık, Daniel A. Lazar, Ramtin Pedarsani, and Dorsa Sadigh. 2021. Emergent Prosociality in Multi-Agent Games Through Gifting. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21*, Zhi-Hua Zhou (Ed.). International Joint Conferences on Artificial Intelligence Organization, 434–442. https://doi.org/10.24963/ijcai.2021/61 Main Track.

[48] Richard Willis and Michael Luck. 2023. Resolving social dilemmas through reward transfer commitments. In *Adaptive and Learning Agents Workshop*. https://kclpure.kcl.ac.uk/ws/portalfiles/portal/207013371/ALA2023_paper_65.pdf

[49] Yuxuan Yi, Ge Li, Yaowei Wang, and Zongqing Lu. 2022. Learning to Share in Networked Multi-Agent Reinforcement Learning. In *Advances in Neural Information Processing Systems*, S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh (Eds.), Vol. 35. Curran Associates, Inc., 15119–15131. https://proceedings.neurips.cc/paper_files/paper/2022/file/61d8577984e4ef0cba20966eb3ef2ed8-Paper-Conference.pdf
[50] Haoqi Zhang and David Parkes. 2008. Value-based policy teaching with active indirect elicitation. In *Proceedings of the 23rd National Conference on Artificial Intelligence - Volume 1* (Chicago, Illinois) *(AAAI'08)*. AAAI Press, 208–214.
[51] Xuezhou Zhang, Yuzhe Ma, Adish Singla, and Xiaojin Zhu. 2020. Adaptive Reward-Poisoning Attacks against Reinforcement Learning. In *Proceedings of the 37th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 119)*, Hal Daumé III and Aarti Singh (Eds.). PMLR, 11225–11234. https://proceedings.mlr.press/v119/zhang20u.html